Zoosyst. Evol. 101 (2) 2025, 955–967 | DOI 10.3897/zse.101.144081

PENSOFT

MUSEUM FÜR NATURKUNDE BERLIN

Substantial mitochondrial gene order rearrangements and differential evolution rates within the family Capitellidae (Annelida)

Xuechun Su[1*], Deyuan Yang[2*], Xiu Wu[1*], Yanan Sun[3], Jian-Wen Qiu[4], Yanjie Zhang[1]

1 *School of Life and Health Sciences, Hainan Province Key Laboratory of One Health, Collaborative Innovation Center of One Health, Hainan University, Haikou, 570228, China*

2 *Institute of Marine Biology, National Taiwan Ocean University, Keelung City, Taiwan*

3 *Laboratory of Marine Organism Taxonomy and Phylogeny, Qingdao Key Laboratory of Marine Biodiversity and Conservation, Institute of Oceanology, Chinese Academy of Sciences, Qingdao, 26000, China*

4 *Department of Biology, Hong Kong Baptist University, Hong Kong, China*

https://zoobank.org/061B2FB2-6875-4A9D-A103-43F4A4A6F066

Corresponding author: Yanjie Zhang (yanjiezhang@hainanu.edu.cn)

Academic editor: Greg Rouse ♦ **Received** 10 December 2024 ♦ **Accepted** 12 April 2025 ♦ **Published** 13 May 2025

Abstract

Capitellidae is a family of marine annelids commonly found in coastal to deep-sea sediments. These annelids are characterized by capillary chaetae at the anterior and long-handled hooks at the posterior part. Although mitochondrial genomes (mtgenomes) are widely used in phylogenetic analyses of invertebrates, their application is limited in many marine annelid families, particularly in Capitellidae. In this study, we obtained complete or nearly complete (except control region) mtgenomes through high-throughput sequencing of eight species across five genera of Capitellidae: *Barantolla* sp., *Capitella teleta*, *Mediomastus* sp., *Notodasus* sp. A, *Notodasus* sp. B, *Notodasus* sp. C, and *Notomastus* sp. A and *Notomastus* sp. B. Our results indicate that species from genera with fewer capillary chaetae (*Barantolla* and *Mediomastus*) exhibit a relatively conserved mitochondrial gene order, while those from other genera show significant gene order rearrangements. Group II intron in *cox1* is found in newly sequenced *Notomastus* sp. B and *Notodasus* sp. A & C. Phylogenetic analysis based on the 13 protein-coding genes (PCGs) or 37 mitochondrial genes (mtgenes) revealed three distinct clades for nine capitellid worms with the mtgenome: Clade 1 includes *Mediomastus* and *Barantolla*; Clade 2 consists of *Notodasus* and *Capitella*; and Clade 3 comprises *Notomastus*. Notably, Clade 2 is sister to Clade 3, and both form the sister group to Clade 1. In contrast, a phylogenetic tree constructed from nuclear genes (ncgenes; *18S*, *28S*, and *H3*) identified *Capitella* as an early branching clade within Capitellidae. The tree based on 37 mtgenes + ncgenes identified the *Capitella* as the sister taxon of *Notodasus* + *Notomastus*. Additionally, the Ka/Ks ratios of 13 PCGs in *Mediomastus* and *Barantolla* were much lower than those in *Notodasus* or *Notomastus*. Together, our results indicate different trajectories of mtgenome evolution in the Capitellidae.

Key Words

Gene order rearrangements, genetic distance, phylogenetics, polychaete

Introduction

Marine annelids are one of the most dominant groups of marine benthic macrofauna (Hutchings 1998), exhibiting remarkable taxonomic diversity with over 11,000 valid species worldwide (Pamungkas et al. 2019). Among these, the Capitellidae comprises free-living marine annelids widely distributed across various marine environments (Rouse et al. 2022). Most capitellids inhabit seafloor substrates, such as contaminated sediments (Tsutsumi 1987), seagrass beds

* These authors contributed equally to this work.

Copyright *Su, X. et al.* This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

(Nakaoka et al. 2002), sandy beaches (Delgado et al. 2003), and some even burrow into squid egg masses (Hartman 1947) or shells of mollusks (Blake 1969).

The type species of *Capitella* (type genus of Capitellidae), *C. capitata*, was originally described as *Lumbricus* by and Fabricius (1780) and Blainville (1828) subsequently erected Capitella for it. In 1862, Grube reviewed the genera *Capitella*, *Dasybranchus*, and *Notomastus*, formally establishing the family Capitellidae (Grube 1862). Carus (1863) further classified these genera into arenicolids, maldanids, and Halelminthea, respectively. Eisig (1887) published a monograph of Capitellidae that included seven genera, and since then, the number of genera in this family has increased substantially. Hartman (1947) illustrated the different distribution of capillary chaetae among 23 genera of Capitellidae known at that time, while Fauchald (1977) constructed a key that included 36 genera, emphasizing the need for further revision within the group. Magalhães and Blake (2022) provided a morphological comparison of 44 genera, documenting 187 species. Currently, 219 accepted species across 41 genera have been recognized, with 32 additional species added and three genera synonymized (Read and Fauchald 2024, Suppl. material 1: table S1). Notably, 21 of these genera are monospecific, while eight genera — *Capitella*, *Dasybranchus*, *Heteromastus*, *Leiochrides*, *Mediomastus*, *Notodasus*, *Notomastus*, and *Scyphoproctus*—contain more than 10 species (Read and Fauchald 2024, Suppl. material 1: table S1).

Genera within Capitellidae are classified based on unique combinations of morphological traits, including the shape of the prostomium and peristomium, the number of thoracic segments, the presence of genital spines, the arrangement of capillaries and hooks, and other features such as sexual phases and the presence or absence of dorsal and lateral grooves or branchiae, and the shape of the pygidium (Green and Gambi 1998; Magalhães and Blake 2022). The distribution of chaetae is particularly crucial for distinguishing genera; however, this trait can change during development, leading to misidentification (Blake et al. 2009; Magalhães and Blake 2022). Furthermore, many capitellid specimens lack posterior parts, and half of the described genera have insufficient information regarding pygidial structures (Magalhães and Blake 2022). Consequently, morphological identification within the family is challenging.

To date, nearly all phylogenetic studies in Capitellidae have relied on various combinations of five genes (i.e., *16S*, *18S*, *28S*, *cox1*, and *H3*). Early studies, which included 3–5 species from 3–5 genera (i.e., *Capitella*, *Notomastus*, *Dasybranchus*, *Barantolla*, and *Heteromastus*), indicated that the genera within Capitellidae are monophyletic (Bleidorn et al. 2003; Struck and Purschke 2005; Rousset et al. 2007; Goto et al. 2013). In Goto (2016), echiura was included in Capitellidae. The most comprehensive molecular phylogeny of Capitellidae, which included 38 species across eight genera, confirmed that *Capitella* is monophyletic, and six genera are not (Tomioka et al. 2018). A recent study based on mitochondrial genes (mtgenes) obtained from Next-Generation Sequencing (NGS) showed that Thalassematidae is sister to Capitellidae (Kobayashi 2023).

The mitochondrial genome (mtgenome) contains vital information regarding gene sequences and orders. It has been used in phylogenetic studies across various marine annelid groups, such as Terebelliformia (Zhong et al. 2008), Aphroditiformia (Zhang et al. 2018), Nereididae (Alves et al. 2020), Serpulidae (Sun et al. 2021), and Chaetopteridae (Wu et al. 2024). Typically, an animal mtgenome comprises one control region and 37 genes, including 13 protein-coding genes (PCGs), 2 rRNAs, and 22 tRNAs. Although the mtgene order of marine annelids has been considered conserved (Jennings and Halanych 2005; Struck et al. 2023), recent studies have revealed exceptions in some groups, including Syllidae (Aguado et al. 2016), certain deep-sea polynoids (Zhang et al. 2018; Hiley et al. 2024), *Ophryotrocha* (Tempestini et al. 2020), and *Hydroides* spp. (Sun et al. 2021). To date, only one mtgenome has been published in Capitellidae (Kobayashi et al. 2022b). The lack of molecular data and the challenges of morphological determination result in difficulties in studying capitellids (Blake et al. 2009). Therefore, more research is needed to elucidate the phylogenetic relationships among different capitellid genera and clarify the delimitation of the morphological characters within Capitellidae.

In China, 23 species in 14 genera of Capitellidae have been recorded, with most found along the southeastern coast (Suppl. material 1: table S2). In this study, the mtgenome and the nuclear *18S*, *28S*, and *H3* genes of eight newly sequenced capitellid species in five genera were added to investigate the phylogenetic relationships in Capitellidae, determine their gene order arrangement patterns, and analyze the genetic substitution rates.

Materials and methods

Sample collection and identification

Samples of Capitellidae collected from China (Shandong, Hainan, Fujian, Guangxi, and Hong Kong) are shown in Fig. 1, Table 1. All specimens were fixed and preserved in 95% ethanol. Identification was carried out following Magalhães and Blake (2022) by the number of thoracic chaetigers, the hooks and capillaries arrangement in the thoracic and abdominal regions, and the presence or absence of achaetous segment or branchiae.

Genome sequencing and quality assessment

In this study, the mtgenomes and nuclear genes (ncgenes, *18S*, *28S*, and *H3*) of eight species of Capitellids were obtained with NGS (Table 1). Several parapodia were used to extract the total DNA using the CTAB method. The DNA was sent to Novogene Co. Ltd. (Tianjin, China) or the Centre for Genomic Sciences (HKU, Hong Kong, China)

Table 1. Collection information of specimens and accession numbers of capitellid mitochondrial genomes, *18S*, *28S*, and *H3* newly obtained in this study.

Species		*Barantolla* sp.	*Capitella teleta*	*Mediomastus* sp.	*Notomastus* sp. A	*Notomastus* sp. B	*Notodasus* sp. A	*Notodasus* sp. B	*Notodasus* sp. C
Sampling site		Yangpudaqiao, Danzhou, Hainan, China	Zhanqiao, Qingdao, Shandong, China	Dayawan, Guangdong, China	Tingkok, Hong Kong, China	Yangpudaqiao, Danzhou, Hainan, China	Tingkok, Hong Kong, China	Changshacun, Danzhou, Hainan, China	Xiamen, Fujian, China
Voucher No.		DZ20240311	SD201108	ZYC17	TK20140629	DZ20240311	TK20140328	DZ20231226	YGLZ220616BYC
Collection date		March 11, 2024	August 2011	January 13, 2022	June 29, 2014	March 11, 2024	May 28, 2014	December 26, 2023	June 16, 2022
Coordinates		19.731°N, 109.194°E	36.06209°N, 120.318405°E	22.774°N, 114.678°E	22.280°N, 114.124°E	19.731°N, 109.194°E	22.280°N, 114.124°E	19.895°N, 109.279°E	24.276°N, 118.175385°E
Mtgenome	**Acc. No.**	PQ010756	PP133665	PP133664	PP133661	PQ010758	PP133663	PQ010757	PP133662
	Length	15,983	17,388	18,972	15,784	17,025	18,725	14,571	17,602
***18S-28S* contig**	**Acc. No.**	PV277003	PV277004	PV277005	PV277006	PV277007	PV277008	PV277009	PV277010
	Length	8,572	8,395	6,592	9,544	7,464	8,041	9,699	7,144
Partial *18S*	**Acc. No.**	PQ365479	PQ365480	PQ365481	PQ365482	PQ365484	PQ365485	PQ365486	PQ365483
	Length	1,672	1,655	1,672	1,687	1,528	1,684	1,684	1,684
Partial *28S*	**Acc. No.**	PQ365487	PQ365488	PQ365489	PQ365490	PQ365492	PQ365493	PQ365494	PQ365491
	Length	621	402	496	750	710	745	746	526
H3	**Acc. No.**	PQ631145	PQ631143	PQ631144	PQ631140	PQ631141	PQ631138	PQ631139	PQ631142
	Length	365	326	304	365	365	365	365	365

for pair-end sequencing on the Illumina HiSeq platform with a read length of 100 bp or 150 bp (Suppl. material 1: table S3). Raw data quality was evaluated using FastQC v0.11.9 under default settings (Brown et al. 2017). Sequence assembly was performed using CLC Genomics Workbench version 7.0.3 (CLCbio, Arhus, Denmark) or NOVOPlasty v2.7.0 (Dierckxsens et al. 2017), SPAdes-3.15.4 (Prjibelski et al. 2020), and GetOrganelle (Jin et al. 2020). The fragments of the mtgenome of *Capitella teleta* were also obtained by PCR and sequenced by first-generation sequencing (FGS). The primers used in PCR are listed in Suppl. material 1: table S4. The fragments were assembled into one contig by CAP3 (Huang and Madan 1999).

Annotation of mtgenome

The contig containing the mtgenome or ncgenes was selected using BLAST V2.13.0 (Camacho and Madden 2013). The mtgenome was annotated using MITOS2 (Bernt et al. 2013b). Since MITOS2 failed to detect the *atp8* in *Notomastus* sp. B, the hmmsearch was used to detect putative *atp8* following Sun et al. (2021). In detail, all possible open reading frames (ORFs) in the forward and reverse strands were predicted using ORFfinder (https://www.ncbi.nlm.nih.gov/orffinder/). A comprehensive database containing *atp8* amino acid sequences from 249 species of Annelida (Suppl. material 1: table S5) was constructed by hmmbuild via HMMER v3.2.1 (Potter et al. 2018). Subsequently, the predicted ORFs (amino acids) were annotated by comparing them with the database using hmmsearch by turning all heuristic filters off (-max) to accept low-scoring matches as positive ones. The boundaries and the start or stop codons of PCGs were manually checked by sequence alignment with that of closely related species. All sequences of the mtgenomes and ncgenes obtained in this study have been deposited in GenBank (Table 1). The GC-skew and AT-skew were calculated according to Perna and Kocher (1995).

Phylogenetic analyses of various datasets

The datasets containing different genes or taxa were used for phylogenetic analysis. A total of 65 taxa were included in the five-gene dataset (i.e., *16S*, *18S*, *28S*, *H3*, and *cox1*) (Suppl. material 1: table S6). For nine capitellids with the mtgenome, i.e., eight new in this study and *Notomastus* sp. LC661358.1 (Kobayashi et al. 2022b), four different datasets of nucleotide sequences were used to determine their evolutionary relationships: 37 mtgenes, 13 PCGs, ncgenes (i.e., *18S*, *28S*, and *H3*), and 37 mtgenes + ncgenes. Opheliidae and Echiuroidea were used as outgroups with Opheliidae as the root. Each gene was aligned using the MAFFT version 7 under default settings (Katoh et al. 2017). Gene concatenation was managed in PhyloSuite (Zhang et al. 2020). The Gblocks Server was used to remove poorly aligned positions or regions by less stringent selection settings (Talavera and Castresana 2007). The best model was detected by PartitionFinder2-2.1.1 in nucleotide mode (Lanfear et al. 2017), and the predicted best model has been listed in Suppl. material 1: table S7. The phylogenetic tree was constructed using both Maximum Likelihood (ML) and Bayesian Inference (BI) methods in PhyloSuite (Zhang et al. 2020), with ML analysis performed using IQ-TREE (Nguyen et al. 2015) with 10,000 standard bootstrap replicates, and BI analysis conducted using MrBayes (Ronquist et al. 2012) with a sampling frequency of 1,000 replicates.

Comparison of gene orders

Pairwise comparisons of mtgene order patterns of both capitellid and outgroups' mtgenomes were performed using CREx2 (Bernt et al. 2007), and the results were used to assess mtgene rearrangements, including transposition (T), reverse transposition (RT), reversal (R), and tandem-duplication-random-loss (TDRL). Because of the

high variability of tRNAs, only PCGs and rRNAs were used to compare the gene arrangement and deduce the hypothetical gene order of their common ancestor.

Base substitutions and nucleotide diversity

Base substitutions were analyzed in 13 PCGs of capitellid mtgenomes. Nucleotide sequences were aligned based on their corresponding amino acid sequences using the MUSCLE module (Edgar 2004) in MEGA11 (Tamura et al. 2021) to ensure accurate codon alignment for the Ka/Ks calculation. The aligned sequences were then used to calculate the Ka (non-synonymous substitutions) to Ks (synonymous substitutions) ratio (Ka/Ks) using KaKs_Calculator 3.0 (Zhang 2022). Ka/Ks > 1 indicates genes undergoing positive selection, Ka/Ks = 1 for neutral evolution, and Ka/Ks < 1 for purifying selection (Yang and Nielsen 2000). Nucleotide diversity of the 13 PCGs and two rRNAs from Capitellidae were detected by DnaSP 6.12.03 (Librado and Rozas 2009) and the gggenes (Wilkins 2023) for visualization.

Data availability

Raw data first reported in this study are available in the NCBI with BioProject No. PRJNA1232974. BioSample accessions are SAMN47255958-SAMN47255963. The raw data of *Notomastus* sp. A and *Notodasus* sp. A are not included since they are corrupted. Please feel free to contact us if anyone needs their assembled contigs. The GenBank accessions of mtgenome, *18S*, *28S*, and *H3* are shown below: PP133660–PP133665 and PQ010756–PQ010758 for mtgenome; PV277003–PV277010 for *18S-28S* contigs; PQ365479–PQ365486 for partial *18S*; PQ365487–PQ365494 for partial *28S*; PQ631138–PQ631145 for *H3*.

Results

General mtgenome features

The mtgenome length of nine capitellids ranged from 14,571 bp to 18,972 bp (Table 1, Suppl. material 1: table S8). Both mtgenomes and PCGs of most of the capitellids are AT-rich (mtgenome: 49.71% to 64.44%; PCGs: 50.05% to 63.96%) (Suppl. material 1: table S8). The AT-skews are -0.282 to 0.081 for mtgenomes and -0.366 to 0.056 for PCGs, respectively; the GC-skews are -0.362 to 0.298 for mtgenomes and -0.422 to 0.298 for PCGs (Suppl. material 1: table S8). For most species, full mtgenomes and PCGs have negative AT-skews and GC-skews, tRNAs have positive AT-skews and GC-skews, and rRNAs have negative GC-skews and positive AT-skews. Only *C. teleta* has negative AT-skews and positive GC-skews for all mtgenes (Fig. 2). ATG or ATA is the start codon for most PCGs. Both complete (TAA or TAG) and incomplete stop codons (TA- and T-) are present (Suppl. material 1: table S9).

Gene order of mtgenome

All the capitellid mtgenomes contain 13 PCGs, 2 rRNA genes, and 22 tRNAs. The mtgenome of *Barantolla* sp. has the same gene order as that of *Mediomastus* sp. except for the location of *trnE*. There is no conserved gene block for the others (Fig. 3). A different number of long non-coding regions (>100 bp)

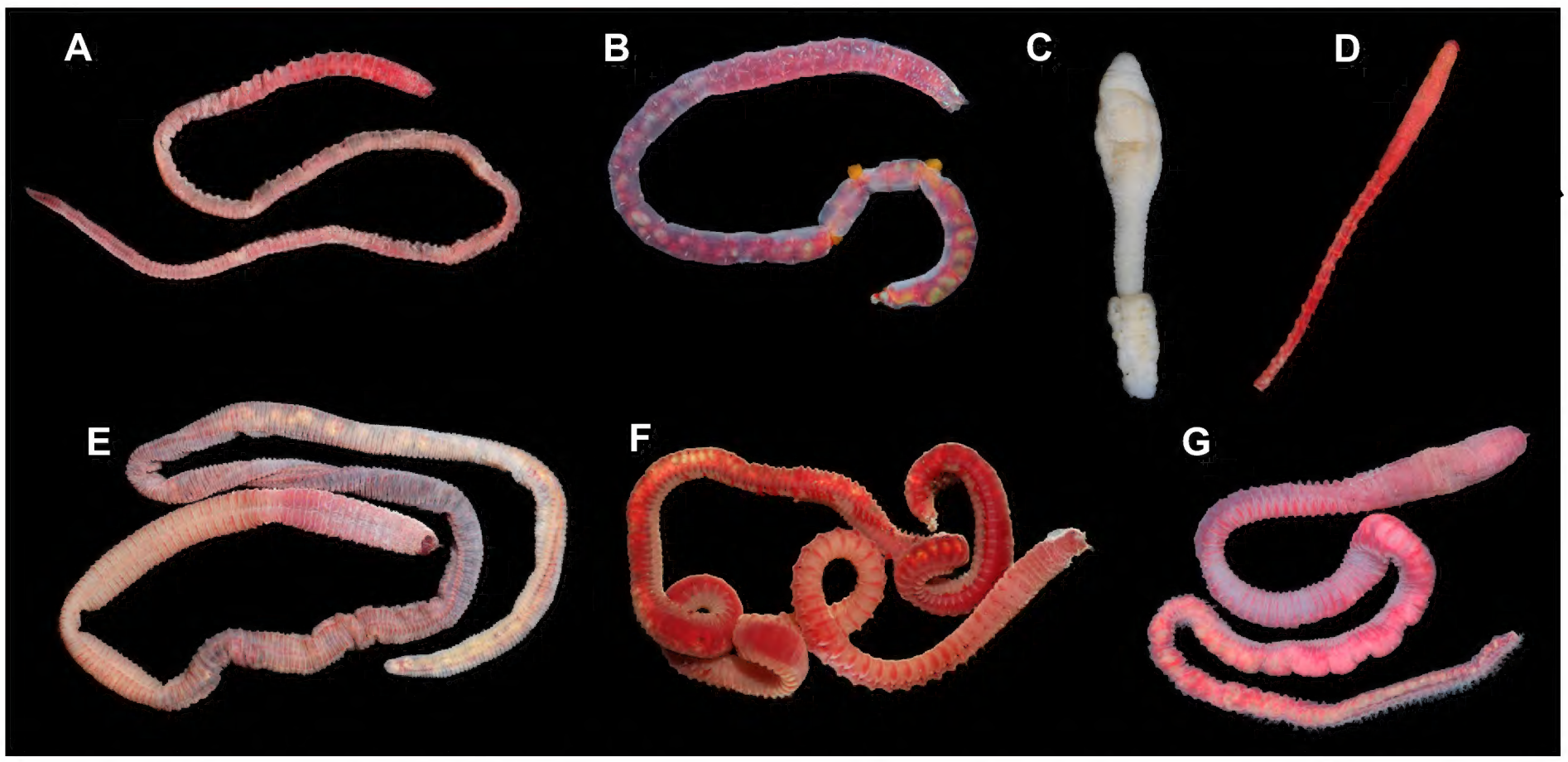

Figure 1. Photos of capitellids newly sequenced in the present study. **A.** *Barantolla* sp.; **B.** *Mediomastus* sp.; **C.** *Notomastus* sp. A; **D.** *Notomastus* sp. B; **E.** *Notodasus* sp. A; **F.** *Notodasus* sp. B; **G.** *Notodasus* sp. C. Figures **C, F** were taken by Su. X.; the others were by Yang D.

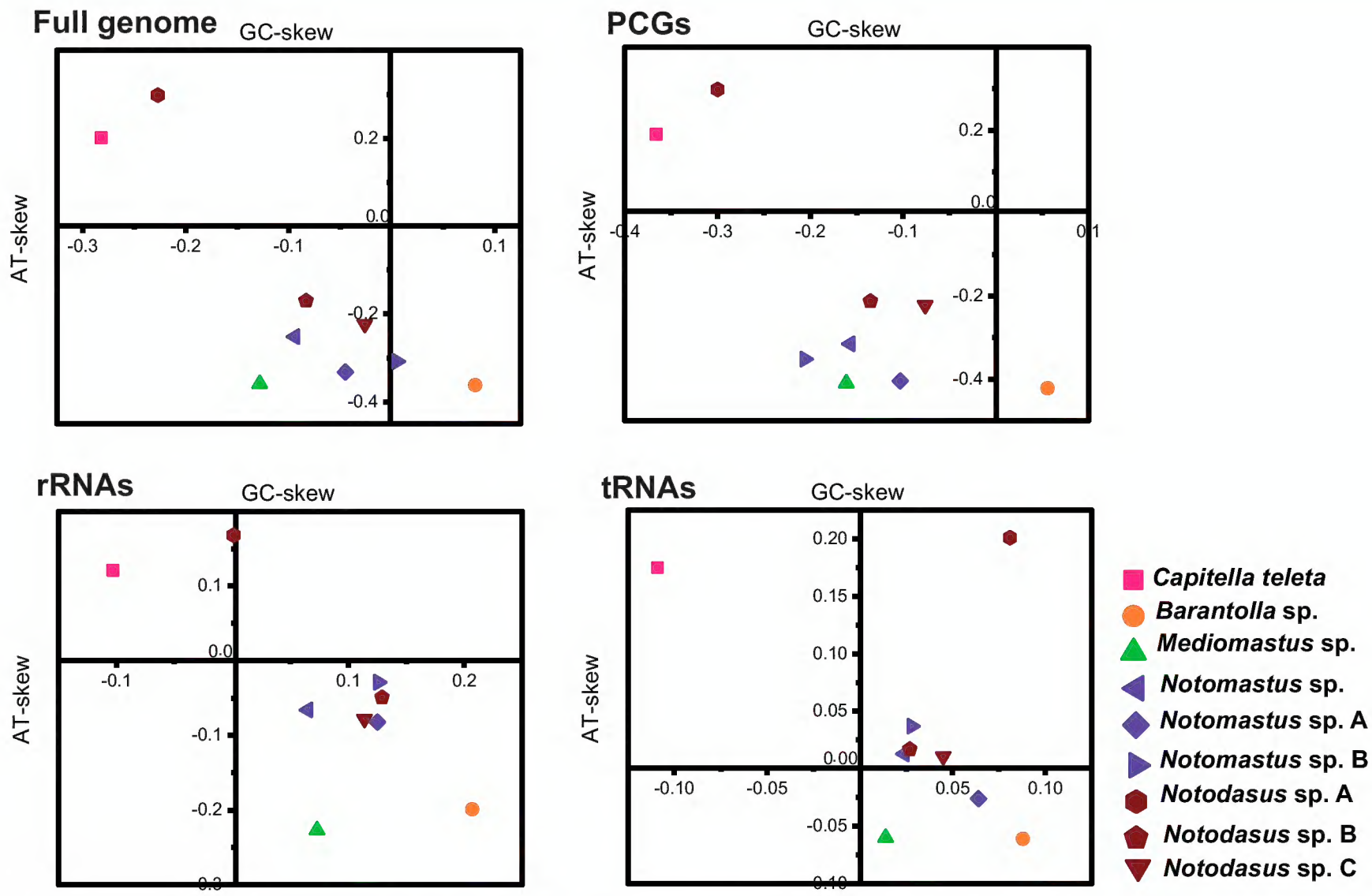

Figure 2. Comparison of AT- and GC-skews of the full genomes, protein-coding genes (PCGs), tRNAs, and rRNAs for the mtgenomes of Capitellidae.

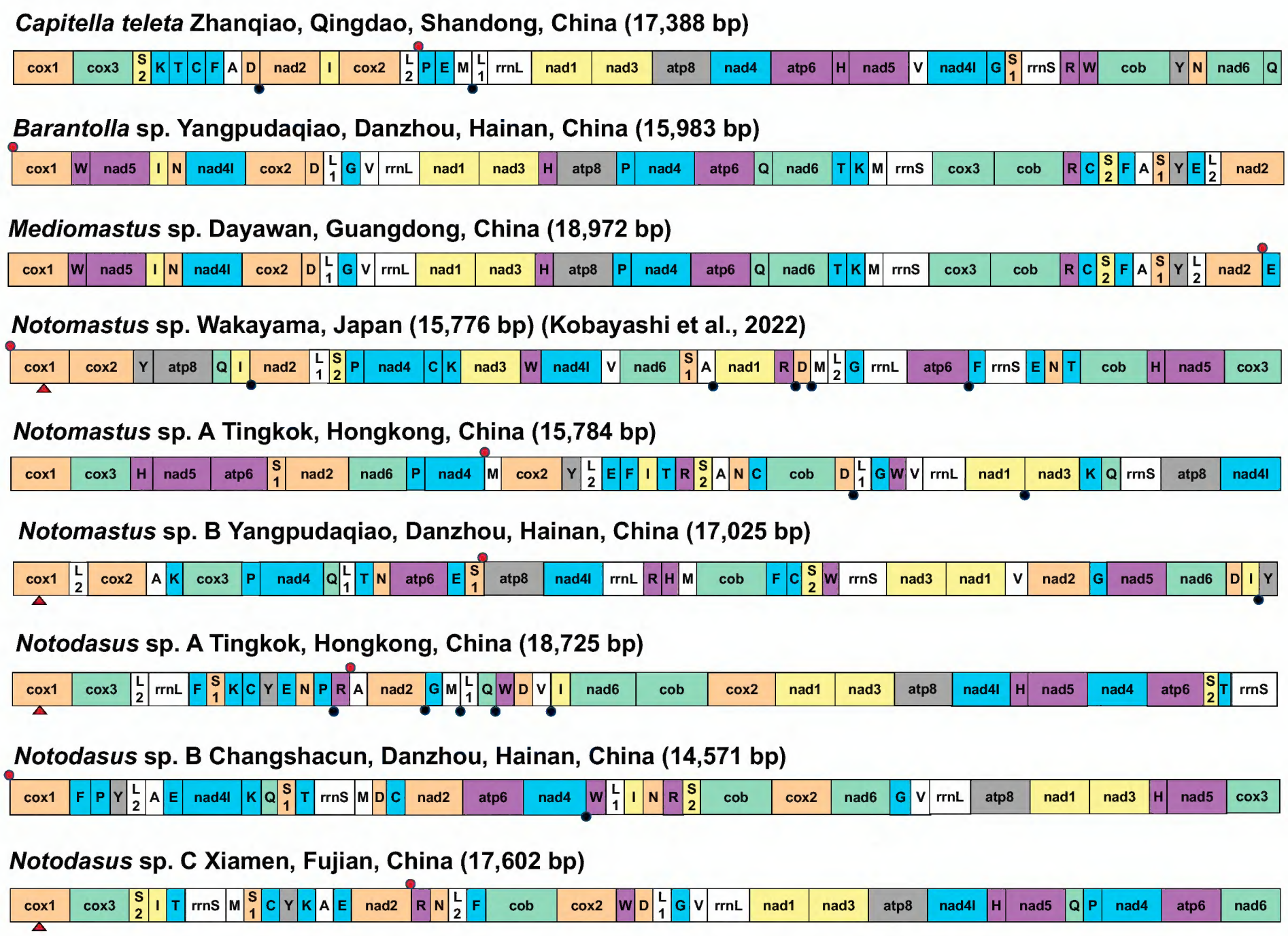
Capitella teleta Zhanqiao, Qingdao, Shandong, China (17,388 bp)
Barantolla sp. Yangpudaqiao, Danzhou, Hainan, China (15,983 bp)
Mediomastus sp. Dayawan, Guangdong, China (18,972 bp)
Notomastus sp. Wakayama, Japan (15,776 bp) (Kobayashi et al., 2022)
Notomastus sp. A Tingkok, Hongkong, China (15,784 bp)
Notomastus sp. B Yangpudaqiao, Danzhou, Hainan, China (17,025 bp)
Notodasus sp. A Tingkok, Hongkong, China (18,725 bp)
Notodasus sp. B Changshacun, Danzhou, Hainan, China (14,571 bp)
Notodasus sp. C Xiamen, Fujian, China (17,602 bp)

Figure 3. Comparison of gene orders of capitellid mtgenomes. Conserved gene clusters in Annelida are indicated by different colors. Black circles represent non-coding regions with > 100 bp sequences between genes. Triangles stand for the long non-coding region (Group II intron) within the *cox1*. Red circles represent the beginning and end of the contig where the putative control region might be located.

between genes of the capitellid mtgenome (Fig. 3, Suppl. material 1: table S10): one in *Notomastus* sp. B and *Notodasus* sp. B, two in *C. teleta* and *Notomastus* sp. A, and five in *Notomastus* sp. and *Notodasus* sp. A. Moreover, a long non-coding region (considered as Group II intron) is in *cox1* of newly sequenced *Notomastus* sp. B and *Notodasus* sp. A & C. Moreover, the putative control region is in different locations of the capitellid mtgenome.

Evolutionary relationships in Capitellidae

The five-gene dataset was used to reconstruct the phylogenetic tree in Capitellidae (Suppl. material 1: table S6). After alignment, the total length of the five genes was 3,257 bp, containing 453 bp for *cox1*, 373 bp for *16S*, 1,473 bp for *18S*, 640 bp for *28S*, and 318 bp for *H3* (Suppl. material 1: table S11). The Capitellidae clade had a high support value for the ML analysis (93%) and a low support value (#) for the BI analysis. Asterisks stand for the highest support values, and number signs stand for the support value < 50 for ML or < 0.5 for BI or not supported. The same applies hereinafter. Capitellidae primarily contained two clades, i.e., Clade A and B. They are lowly supported by the ML and BI analysis (A, ML/BI: 64/#; B, ML/BI: 54/#). Clade A contains *Barantolla*, *Decamastus*, *Heteromastus*, *Leiochrides*, *Mediomastus*, *Mastobranchus*, and *Notomastus*; Clade B covers species from all genera included except *Mastobranchus*. *Capitella* is monophyletic (Fig. 4). Seven genera (i.e., *Barantolla*, *Decamastus*, *Heteromastus*, *Leiochrides*, *Mediomastus*, *Notomastus*, *Notodasus*) are paraphyletic or polyphyletic. Notably, the support values are only moderate for most nodes in the tree.

Relationships of nine capitellids with mtgenome in phylogenetic trees based on 37 mtgenes and 13 PCGs are identical, while the location of *C. teleta* is different in the tree based on 37 mtgenes + ncgenes (Fig. 5). In all the trees, the clade Capitellidae is well supported by both the ML and BI analyses (ML/BI: */*). Two clades are present within Capitellidae: *Mediomastus* and *Barantolla* consistently clustered together with fully supported (ML/BI: */*) and form Clade 1, and species of *Notomastus*, *Notodasus*, and *Capitella* form Clade 2 with a high support value (ML/BI: */* in all the trees) (Fig. 5). *Notomastus* sp. A + (*Notomastus* sp. + *Notomastus* sp. B) form Clade 2a, and (*C. teleta* + *Notodasus* sp. A) + (*Notodasus* sp. B + *Notodasus* sp. C) form another clade for both 37 mtgenes and 13 PCGs (Fig. 5A). When nuclear genes are included, the location of *C. teleta* is different, and *C. teleta* + (*Notodasus* spp. + *Notomastus* spp.) from Clade 2 (Fig. 5B). Moreover, the support values of clades with *C. teleta* are low in the 37-mtgenes tree, and the support values increase when nuclear genes are added.

Rearrangements of mtgene

The evolutionary patterns of PCGs and rRNA are shown in Fig. 6. A side-by-side comparison of the phylogenetic tree and gene order by TreeREx revealed the evolutionary pattern of PCGs and rRNAs (Fig. 6). The gene cluster *cox1-cox2-atp8* is present in the hypothetical ancestor of Capitellidae & *Urechis*, but not in Capitellidae only. One transposition (*and5-nad4l*) happened from the capitellid ancestor to the common ancestor of *Mediomastus* and *Barantolla*, and three transpositions (*nad6*, *cox3*, *nad2*) from the capitellid ancestor to the common ancestor of *Capitella*, *Notomastus*, and *Notodasus* (Fig. 6, Suppl. material 1: fig. S1). A particularly complex process was detected from the common ancestor of *Notomastus* and *Notodasus* to the present species, and more than half of the mtgenes are involved in the rearrangement scenario (Fig. 6, Suppl. material 1: fig. S1).

Purifying selection of 13 PCGs and nucleotide diversity (Pi)

The Ka/Ks of the 13 PCGs of Capitellidae ranged from 0.031 (*cox1*) to 0.385 (*atp8*), suggesting that they have undergone strong purification selection (Fig. 7A). When comparing the Ka/Ks of PCGs in different genera, the relatively conserved *Mediomastus* & *Barantolla* revealed much lower Ka/Ks values for each gene than *Notodasus* or *Notomastus* (Fig. 7B). Nucleotide diversity (Pi) analyses of 13 PCGs and 2 rRNAs from capitellid mtgenomes showed that the average nucleotide diversity of *Barantolla* & *Mediomastus* was lower than that of *Notodasus* and *Notomastus* (Fig. 7C). The Pi values of *rrnS* and *atp6* are low (< 0.1). The Pi values of *nad5*, *nad1*, and *nad4* are high (>0.5) (Fig. 7C).

Discussion

Structural features and mtgene rearrangement of capitellid mtgenomes

Traditionally, mtgenomes in marine annelids were thought to be compact. However, recent studies have identified long non-coding regions in Siboglinidae (Li et al. 2015), *Owenia fusiformis* (Weigert et al. 2016), Syllidae (Aguado et al. 2016), Serpulidae (Seixas et al. 2017; Sun et al. 2021), *Decemunciger* (Bernardino et al. 2017), certain deep-sea species of the Polynoidae (Zhang et al. 2018; Hiley et al. 2024), *Armandia* (Kobayashi et al. 2022b), and *Synelmis* (Huč et al. 2024). In this study, the gene orders in *Barantolla* and *Mediomastus* are identical (except *trnE*), but there is substantial variability in *Capitella*, *Notomastus*, and *Notodasus* (Fig. 3). In contrast, long non-coding regions (>100 bp) were not found in *Barantolla* and *Mediomastus* and were found in species of *Capitella*, *Notomastus* spp., and *Notodasus* spp. (excluding *Notodasus* sp. C), suggesting that these regions might be involved in mtgene rearrangement and have once been present and removed afterward in *Notodasus* sp. C. This implies that different strategies or mechanisms may govern the rearrangement of capitellid mtgenomes.

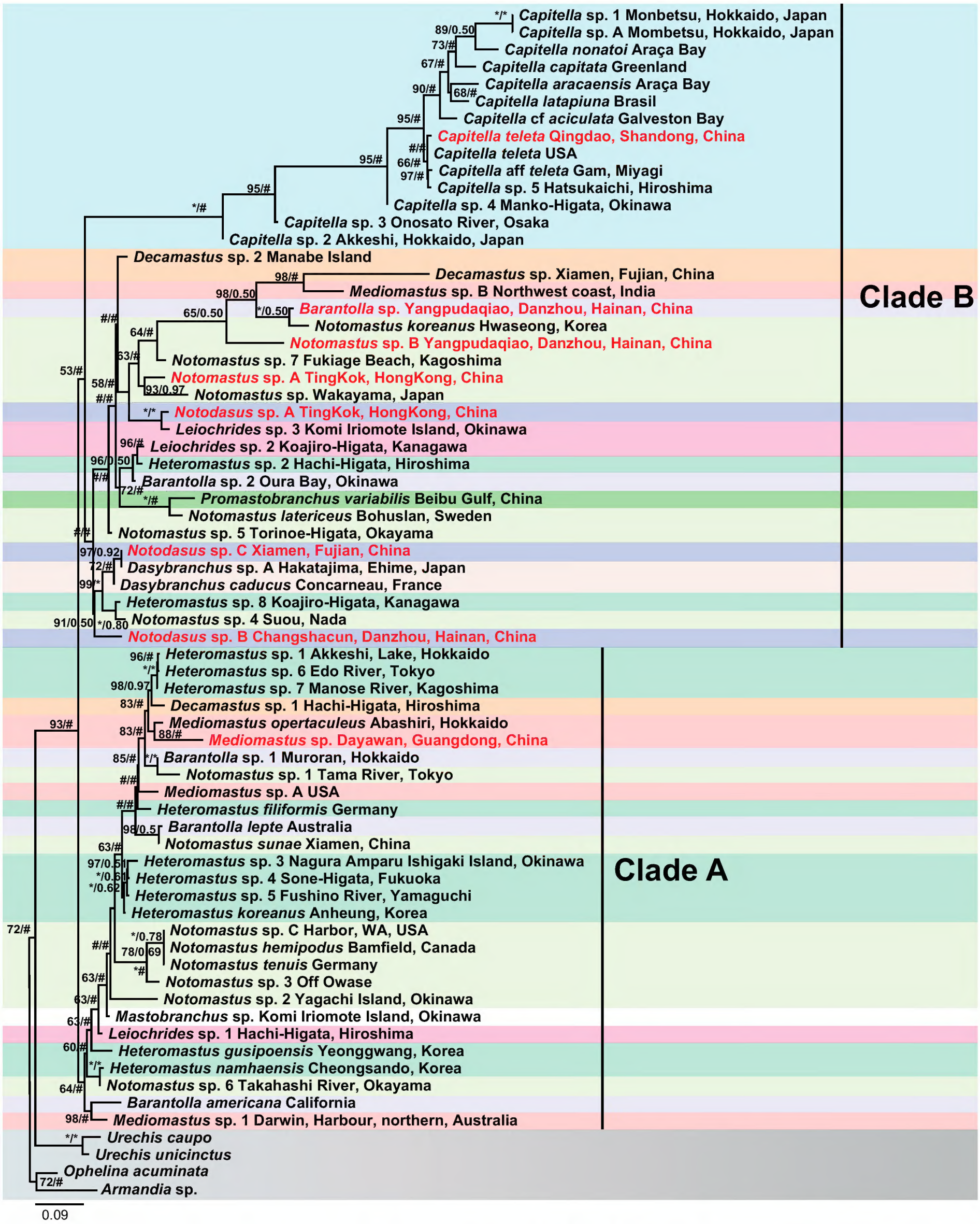
Capitella sp. 1 Monbetsu, Hokkaido, Japan
Capitella sp. A Mombetsu, Hokkaido, Japan
Capitella nonatoi Araça Bay
Capitella capitata Greenland
Capitella aracaensis Araça Bay
Capitella latapiuna Brasil
Capitella cf aciculata Galveston Bay
Capitella teleta Qingdao, Shandong, China
Capitella teleta USA
Capitella aff teleta Gam, Miyagi
Capitella sp. 5 Hatsukaichi, Hiroshima
Capitella sp. 4 Manko-Higata, Okinawa
Capitella sp. 3 Onosato River, Osaka
Capitella sp. 2 Akkeshi, Hokkaido, Japan
Decamastus sp. 2 Manabe Island
Decamastus sp. Xiamen, Fujian, China
Mediomastus sp. B Northwest coast, India
Barantolla sp. Yangpudaqiao, Danzhou, Hainan, China
Notomastus koreanus Hwaseong, Korea
Notomastus sp. B Yangpudaqiao, Danzhou, Hainan, China
Notomastus sp. 7 Fukiage Beach, Kagoshima
Notomastus sp. A TingKok, HongKong, China
Notomastus sp. Wakayama, Japan
Notodasus sp. A TingKok, HongKong, China
Leiochrides sp. 3 Komi Iriomote Island, Okinawa
Leiochrides sp. 2 Koajiro-Higata, Kanagawa
Heteromastus sp. 2 Hachi-Higata, Hiroshima
Barantolla sp. 2 Oura Bay, Okinawa
Promastobranchus variabilis Beibu Gulf, China
Notomastus latericeus Bohuslan, Sweden
Notomastus sp. 5 Torinoe-Higata, Okayama
Notodasus sp. C Xiamen, Fujian, China
Dasybranchus sp. A Hakatajima, Ehime, Japan
Dasybranchus caducus Concarneau, France
Heteromastus sp. 8 Koajiro-Higata, Kanagawa
Notomastus sp. 4 Suou, Nada
Notodasus sp. B Changshacun, Danzhou, Hainan, China
Heteromastus sp. 1 Akkeshi, Lake, Hokkaido
Heteromastus sp. 6 Edo River, Tokyo
Heteromastus sp. 7 Manose River, Kagoshima
Decamastus sp. 1 Hachi-Higata, Hiroshima
Mediomastus opertaculeus Abashiri, Hokkaido
Mediomastus sp. Dayawan, Guangdong, China
Barantolla sp. 1 Muroran, Hokkaido
Notomastus sp. 1 Tama River, Tokyo
Mediomastus sp. A USA
Heteromastus filiformis Germany
Barantolla lepte Australia
Notomastus sunae Xiamen, China
Heteromastus sp. 3 Nagura Amparu Ishigaki Island, Okinawa
Heteromastus sp. 4 Sone-Higata, Fukuoka
Heteromastus sp. 5 Fushino River, Yamaguchi
Heteromastus koreanus Anheung, Korea
Notomastus sp. C Harbor, WA, USA
Notomastus hemipodus Bamfield, Canada
Notomastus tenuis Germany
Notomastus sp. 3 Off Owase
Notomastus sp. 2 Yagachi Island, Okinawa
Mastobranchus sp. Komi Iriomote Island, Okinawa
Leiochrides sp. 1 Hachi-Higata, Hiroshima
Heteromastus gusipoensis Yeonggwang, Korea
Heteromastus namhaensis Cheongsando, Korea
Notomastus sp. 6 Takahashi River, Okayama
Barantolla americana California
Mediomastus sp. 1 Darwin, Harbour, northern, Australia
Urechis caupo
Urechis unicinctus
Ophelina acuminata
Armandia sp.
Clade B
Clade A
0.09

Figure 4. Phylogenetic tree of Capitellidae based on concatenated data of *16S*, *18S*, *28S*, *H3*, and *cox1*. The tree shown is from ML analysis. The species newly sequenced for this study are in red. Bootstrap support values of ML (left) and posterior probability of BI (right) are indicated above the nodes. Asterisks stand for the highest support values, and number signs stand for the support value < 50 for ML or < 0.5 for BI or not supported.

The increased substitution rate may lead to a high rate of gene rearrangement (Shao et al. 2003). Our results indicate that the substitution rate of capitellid mtgenes varies. In species with relatively conserved gene orders, such as *Barantolla* and *Mediomastus*, the substitution rates for each PCG are significantly lower than those in *Notomastus* and *Notodasus* (Fig. 7B). This suggests that greater variability in mtgene order may correspond to higher substitution rates, as previous studies showed (Bernt et al. 2013a; Shao et al. 2003).

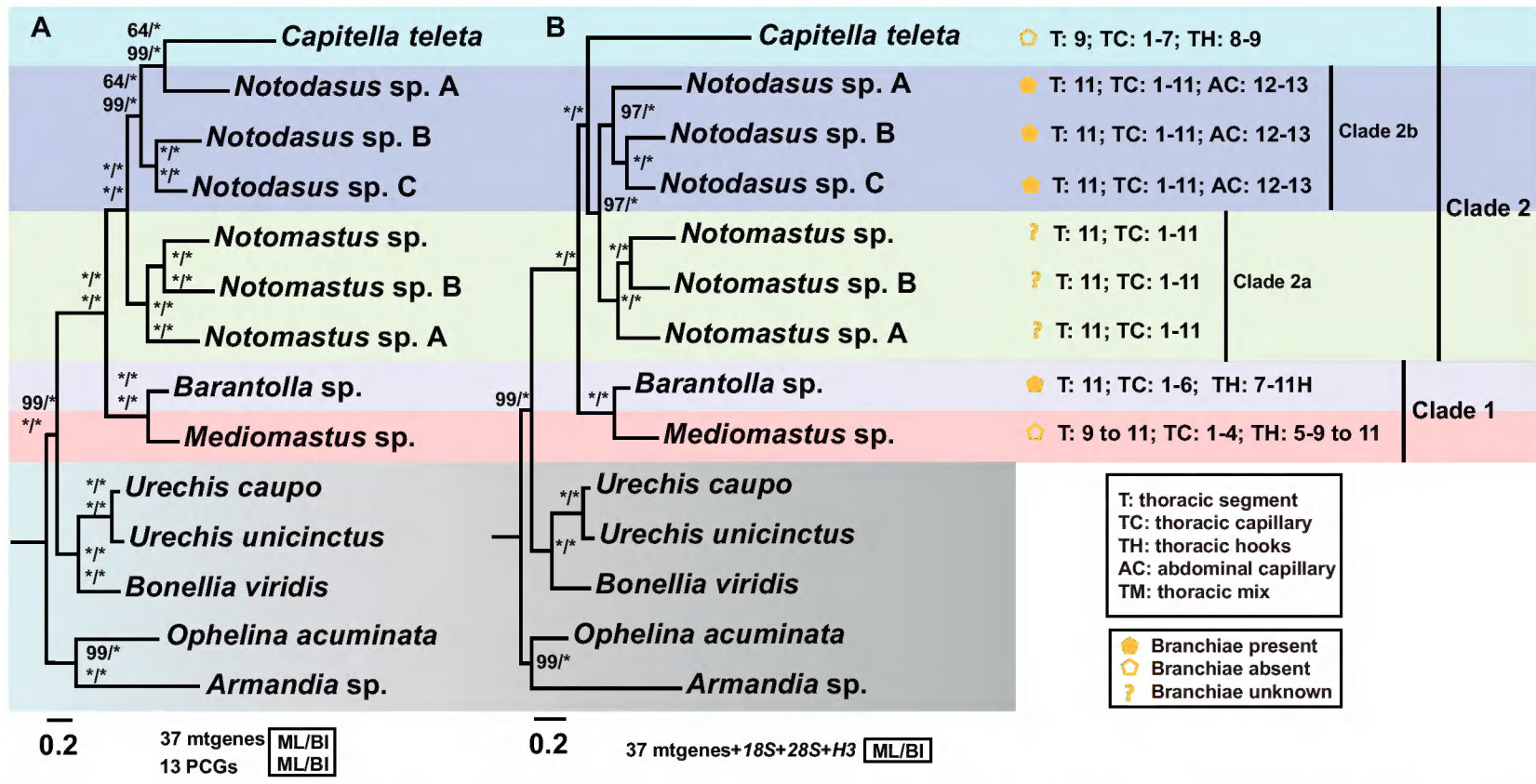

Figure 5. The phylogenetic tree of Capitellidae based on 37 mtgenes, 13 PCGs (**A**), or 37 mtgenes + ncgenes (*18S* + *28S* + *H3*) (**B**). The tree shown is from ML analysis based on 37 mtgenes + ncgenes (*18S* + *28S* + *H3*) and 37 mtgenes. Asterisks stand for the highest support values, and number signs stand for the support value < 50 for ML or < 0.5 for BI or not supported.

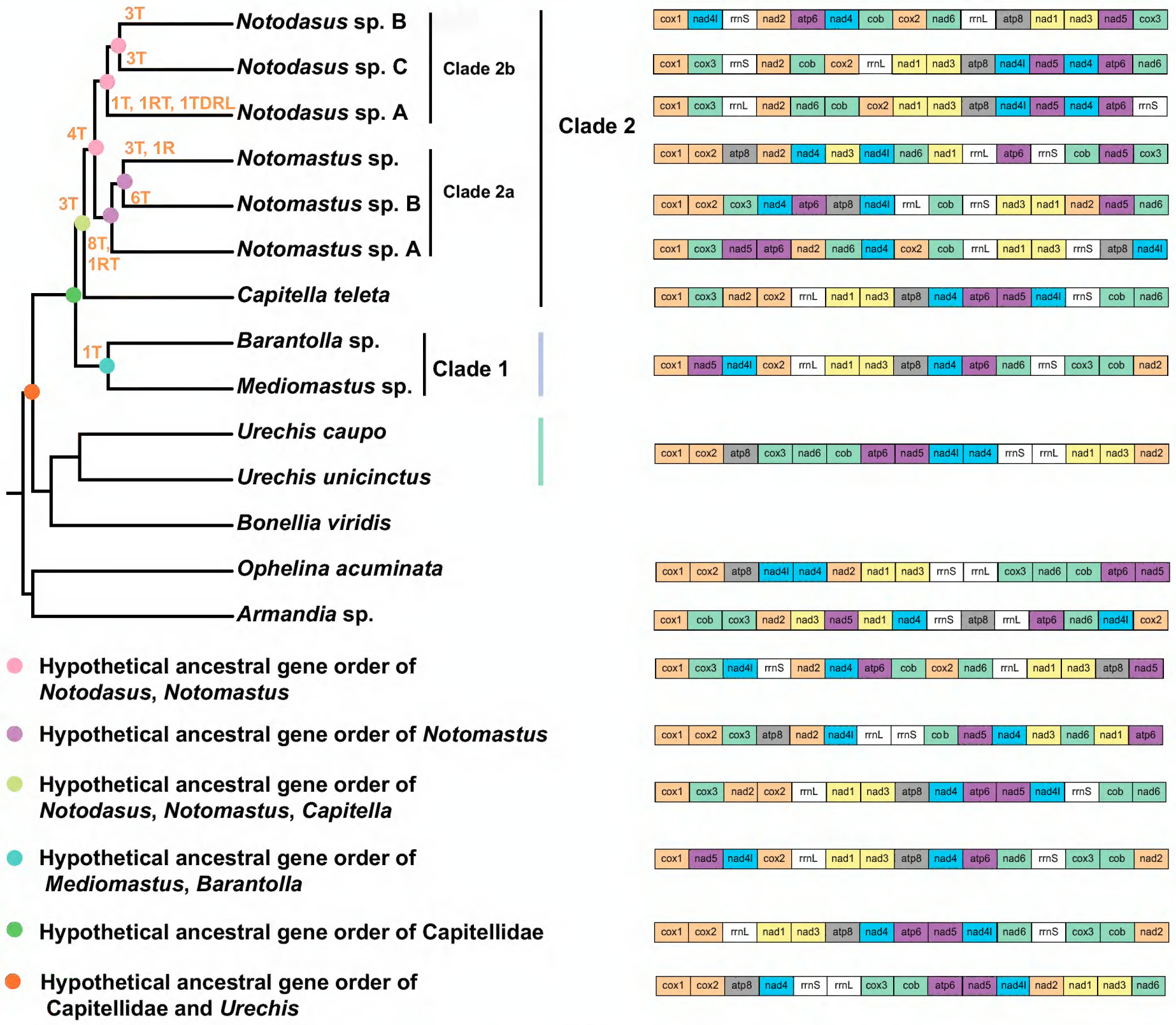

Figure 6. Possible mechanisms of mtgene rearrangement in Capitellidae estimated using TreeREx. Conserved gene clusters in Annelida are indicated by different colors. Only PCGs and rRNAs are included. R, reversal; T, transposition; RT, reverse transposition; TDRL, tandem-duplication-random-loss.

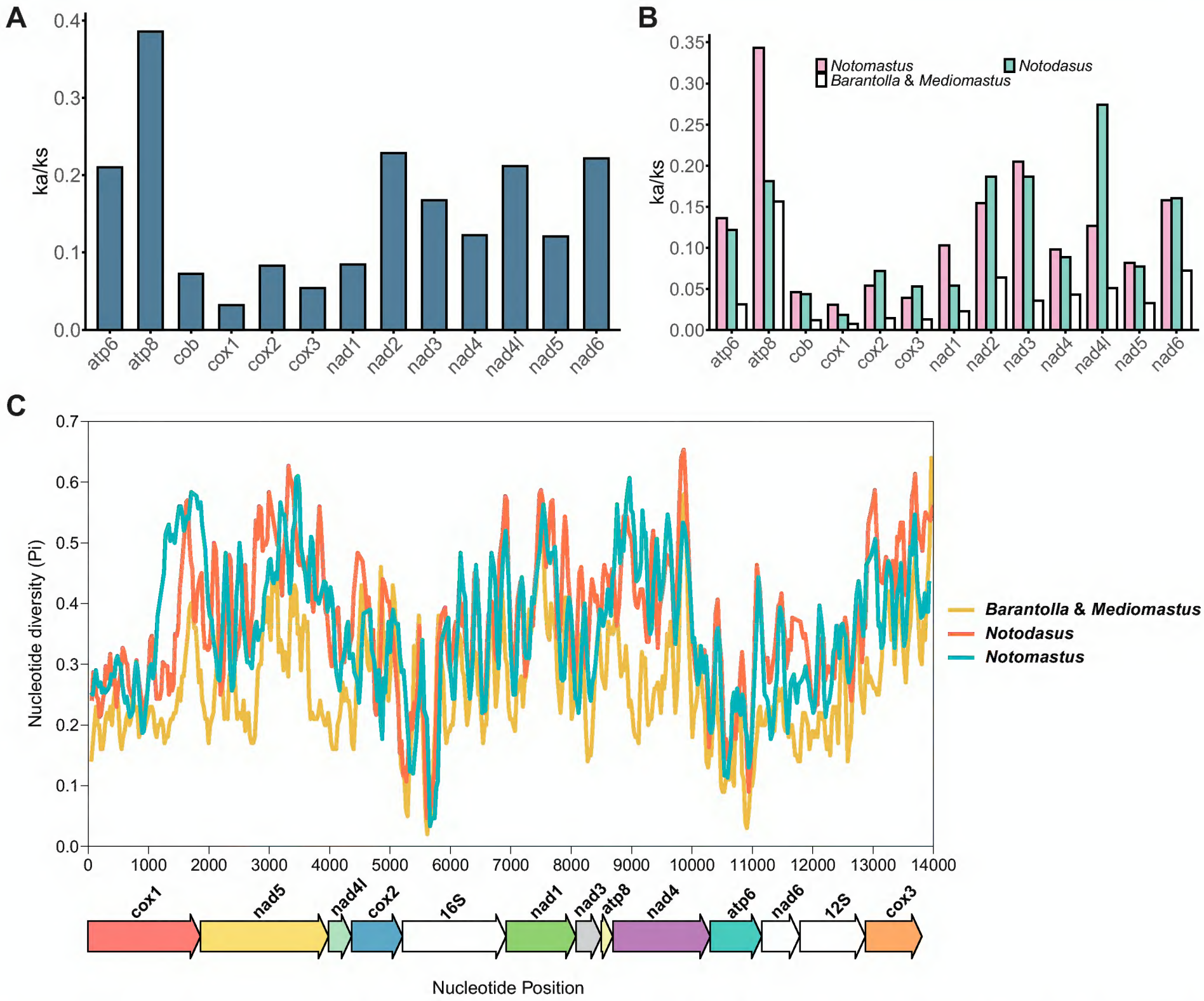

Figure 7. Base substitution rate of the 13 mitochondrial PCGs in Capitellidae. **A.** Ka/Ks for each mtgene; **B.** Ka/Ks of *Notodasus* and *Notomastus* and *Mediomastus* & *Barantolla*. Ka/Ks < 1 indicates purification selection. (C) Nucleotide diversity analysis of 13 PCGs + two rRNAs based on Capitellidae. The yellow line represents *Barantolla* & *Mediomastus*, the red line represents *Notodasus*, and the blue line represents *Notomastus*.

The gene order in the nine capitellid mtgenomes differs significantly from that of other annelids, and most conserved gene clusters in Annelida were split in capitellids (Jennings and Halanych 2005; Struck et al. 2023). Also, they exhibit intragenus variability. Previous studies showed that Echiura is the sister taxon to Capitellidae (Tomioka et al. 2018; Kobayashi 2023). The conserved gene cluster *cox1-cox2-atp8* is present in both *Urechis* spp. and *Notomastus* sp. (Tomioka et al. 2018; Kobayashi et al. 2022b), which supports their close relationship, but the similarity of mtgene order between Echiura and Capitellidae was concluded on a limited sampling size. In our results, *cox1-cox2-atp8* is not present in the hypothetical ancestral mtgene order of Capitellidae (Fig. 6), indicating the evolution of mtgene order in Capitellidae is more complex than previously thought. Further work needs to be done on more taxa covering more genera to more accurately predict the ancestral mtgene order in Capitellidae.

Mtgenome and ncgene could be combined to determine the phylogenetic relationships within Capitellidae

In recent years, phylogenetic analyses of the Capitellidae, based on limited molecular data and taxa, have produced conflicting results (Rousset et al. 2007; Goto et al. 2013). The most comprehensive phylogenetic analysis of Capitellidae utilized a concatenated gene marker dataset, including *18S*, *28S*, *H3*, and *cox1* (Tomioka et al. 2018). Our findings align with this earlier work, confirming that *Capitella* is monophyletic while most of the other examined genera are paraphyletic or polyphyletic. Moreover, the topology of the tree changed when different outgroups were included for the five-gene tree (Fig. 4, Suppl. material 1: fig. S2). This discrepancy may arise from the under-sampling of taxonomic units or the use of limited genetic markers since there were only one or two markers for nearly half of the taxa (40% in this study for Fig. 4, Suppl. material 1: table S6).

Additionally, the accuracy of species identification by morphology is crucial, as this will affect our interpretation of molecular phylogenetic analyses. Due to its low mutation rates, the *18S* is not ideal for distinguishing closely related species in Capitellidae, particularly for *Dasybranchus*: there is only partial *18S* for *D. caducus* and partial *18S* and *28S* for *Dasybranchus* sp. The *18S* of both *D. caducus* and *Dasybranchus* sp. have a high sequence identity compared to those of three *Notodasus* spp. in this study, 99.1–100% and 98.9–99.8%, respectively (Suppl. material 1: table S12). This high level of conservation in *18S* suggests that this gene lacks the resolution for species-level delimitation. In contrast, for *28S*, *Dasybranchus* sp. showed 99.6% identity with *Notodasus* sp. C and 93% identity with *Notodasus* sp. B (Suppl. material 1: table S12). More evidence is needed to determine whether the *D. caducus* in Fig. 4 is the same species as *Notodasus* sp. C.

Mtgenomes are increasingly employed in phylogenetic studies due to the wealth of genotypic information they provide (Boore 1999). They have a higher likelihood of accurately tracking short internodes compared to ncgenes (Moore 1995). The tree topology of Capitellidae presented in this paper, based on datasets that include ncgenes, is different from trees constructed with mtgenes only (Fig. 5). This indicates that mtgenomes and ncgenes might have different evolution trajectories, and they together could contribute to elucidating the complex relationships within Capitellidae. Our results with the combined mtgenes and ncgenes showed that *Barantolla* is the sister group to *Mediomastus*, with *Notomastus* and *Notodasus* being monophyletic. Notably, the phylogenetic position of *Capitella* is different in the tree based on ncgenes, mtgenes, and mtgenes + ncgenes (Fig. 5, Suppl. material 1: fig. S3). It is important to note that the phylogenetic placement of *Capitella* remains unclear. When mitochondrial rRNA and tRNA genes are included in the phylogenetic dataset, the clade that includes *Capitella* generally shows lower support values (Fig. 5A). In trees constructed using ncgenes only, *Capitella* appears as the sister taxon to all the other capitellids (Suppl. material 1: fig. S3). Notably, their phylogenetic tree support values were not high (59/0.84 for ncgenes; 69/0.76 for *18S-28S*), and *Capitella* might be misplaced (Suppl. material 1: fig. S3). This might be due to the difficulty in the alignment of the ribosomal genes. The tree topology might be greatly affected by long-branch attraction. Phylogenetic trees based on *nad2*, *nad4*, and *nad6* exhibit the same topology as that based on the mtgenome, while *atp6*, *cob*, *nad5*, and *rrnL* exhibit the same topology as that based on the mtgenes + ncgenes (Suppl. material 1: fig. S4). When *Capitella* was not considered, most trees based on a single gene supported a close relationship between *Barantolla* and *Mediomastus* and the monophyly of *Notomastus* and *Notodasus*. It is worth noting that our results are based on a limited taxon sampling when considering there are 41 genera in Capitellidae. Incorporating more taxa into future studies is essential to better elucidate the relationships among these genera.

Group II Introns in cox1 of capitellid mtgenomes

Since the first report of Group II introns in *Nephtys* (Vallès et al. 2008), they have been found in different taxa of polychaete, such as Glyceridae (Richter et al. 2015), *Decemunciger* (Bernardino et al. 2017), Travisiidae (Kobayashi et al. 2022a), *Notomastus* and *Armandia* (Kobayashi et al. 2022b), and Chaetopteridae (Wu et al. 2024). The length of Group II intron varies in different species: 1,819 in *Nephty* sp. (Vallès et al. 2008), 1,647 in *Decemunciger* sp. (Bernardino et al. 2017), 1,646 in *Mesochaetopterus tingkokensis*, and 365 bp in *Phyllochaetopterus* sp. (Wu et al. 2024). Group II intron in the *cox1* gene is in the newly sequenced *Notomastus* sp. B, *Notodasus* sp. A & C. Its length in these capitellids ranges from 273 to 447 bp, and its insertion position into the *cox1* genes is the same. This result indicates that the Group II intron might be present in the common ancestor of *Notomastus* and *Notodasus*. However, answering whether they are all shorter than those of some other polychaete species requires the sequencing of more capitellids.

Conclusions

In this study, we obtained complete or nearly complete (except control region) mtgenomes through high-throughput sequencing of eight species across five genera of Capitellidae. Our results indicate that species from genera with fewer capillary chaetae (*Barantolla* and *Mediomastus*) exhibit a relatively conserved mitochondrial gene order, while those from other genera show significant gene order rearrangements. Phylogenetic analysis supported the close relationships of *Mediomastus* and *Barantolla*, *Notodasus* and *Notomastus*, but the location of *Capitella* is undetermined. Additionally, the Ka/Ks ratios of 13 PCGs in *Mediomastus* and *Barantolla* were much lower than those in *Notodasus* or *Notomastus*. Together, our results indicate different trajectories of mtgenome evolution in the Capitellidae.

Acknowledgments

This work was financially supported by the Hainan Province Science and Technology Special Fund (Grant No. ZDYF2024SHFZ101), the National Natural Science Foundation of China (Grant Nos. 42306132 and 42106118), the Shandong Provincial Natural Science Foundation (Grant No. 2023HWYQ-101), the Taishan Scholars Program (Grant No. tsqn:202306288), and the Hainan University Research Start-up Fund Project (Grant No. KYQD(ZR)-21144).

References

Aguado MT, Richter S, Sontowski R, Golombek A, Struck TH, Bleidorn C (2016) Syllidae mitochondrial gene order is unusually variable for Annelida. Gene 594(1): 89–96. https://doi.org/10.1016/j.gene.2016.08.050

Alves PR, Halanych KM, Santos CSG (2020) The phylogeny of Nereididae (Annelida) based on mitochondrial genomes. Zoologica Scripta 49(3): 366–378. https://doi.org/10.1111/zsc.12413

Bernardino AF, Li Y, Smith CR, Halanych KM (2017) Multiple introns in a deep-sea Annelid (*Decemunciger*, Ampharetidae) mitochondrial genome. Scientific reports 7(1): 4295. https://doi.org/10.1038/s41598-017-04094-w

Bernt M, Merkle D, Ramsch K, Fritzsch G, Perseke M, Bernhard D, Schlegel M, Stadler PF, Middendorf M (2007) CREx: inferring genomic rearrangements based on common intervals. Bioinformatics 23(21): 2957–2958. https://doi.org/10.1093/bioinformatics/btm468

Bernt M, Braband A, Schierwater B, Stadler PF (2013a) Genetic aspects of mitochondrial genome evolution. Molecular Phylogenetics and Evolution 69: 328–338. https://doi.org/10.1016/j.ympev.2012.10.020

Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritzsch G, Pütz J, Middendorf M, Stadler PF (2013b) MITOS: Improved de novo metazoan mitochondrial genome annotation. Molecular Phylogenetics and Evolution 69: 313–319. https://doi.org/10.1016/j.ympev.2012.08.023

Blainville HD (1828) Vers a sang rouge. In: Levrault FG (Ed.) Dictionnaire des Sciences Naturelles, Strasbourg, 368–501.

Blake JA (1969) Systematics and ecology of shell-boring polychaetes from New England. American Zoologist 9(3): 813–820. https://doi.org/10.1093/icb/9.3.813

Blake JA, Grassle JP, Eckelbarger KJ (2009) *Capitella teleta*, a new species designation for the opportunistic and experimental Capitella sp. I, with a review of the literature for confirmed records. Zoosymposia 2: 25–53. https://doi.org/10.11646/zoosymposia.2.1.6

Bleidorn C, Vogt L, Bartolomaeus T (2003) New insights into polychaete phylogeny (Annelida) inferred from 18S rDNA sequences. Molecular phylogenetics and evolution 29(2): 279–288. https://doi.org/10.1016/S1055-7903(03)00107-6

Boore JL (1999) Animal mitochondrial genomes. Nucleic Acids Research 27(8): 1767–1780. https://doi.org/10.1093/nar/27.8.1767

Brown J, Pirrung M, McCue LA (2017) FQC Dashboard: integrates FastQC results into a web-based, interactive, and extensible FASTQ quality control tool. Bioinformatics 33(19): 3137–3139. https://doi.org/10.1093/bioinformatics/btx373

Camacho C, Madden T (2013) "BLAST+ Release Notes. Mar 12 [Updated 2023 Apr 24]," in BLAST® Help (National Center for Biotechnology Information (US). 2008-, Bethesda (MD). https://www.ncbi.nlm.nih.gov/books/NBK131777/

Carus JV (1863) Vermes. In: Peters WCH, Carus JV, Gerstäcker CEA (Eds) Handbuch der zoologie: Leipzig, Wilhelm Engelmann, 422–484. http://biodiversitylibrary.org/page/1092677

Delgado JD, Núñez J, Riera R, Monterroso O (2003) Abundance and diversity patterns of annelids from intertidal sandy beaches in Iceland. Hydrobiologia 496: 311–319. https://doi.org/10.1023/A:1026121604617

Dierckxsens N, Mardulyn P, Smits G (2017) NOVOPlasty: *de novo* assembly of organelle genomes from whole genome data. Nucleic Acids Research 45: e18. https://doi.org/10.1093/nar/gkw955

Edgar RC (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research 32(5): 1792–1797. https://doi.org/10.1093/nar/gkh340

Eisig H (1887) Monographie der Capitelliden des Golfes von Neapel und der angrenzenden Meeres-abschnitte: Atlas. Verlag Von R. Friedländer & Sohn, Berlin, Friedländer, 906pp

Fabricius O (1780) Fauna Groenlandica: systematice sistens animala Groenlandiae occidentalis hactenus indagata, quod nomen specificium. Hafniae et Lipsiae, Copenhagen, Denmark XVi: 452. https://doi.org/10.5962/bhl.title.13489

Fauchald K (1977) The polychaete worms. Definitions and keys to the orders, families and genera. Natural History Museum of Los Angeles County, Science Series, 188 pp.

Goto R (2016) A comprehensive molecular phylogeny of spoon worms (Echiura, Annelida): Implications for morphological evolution, the origin of dwarf males, and habitat shifts. Molecular Phylogenetics and Evolution 99: 247–260. https://doi.org/10.1016/j.ympev.2016.03.003

Goto R, Okamoto T, Ishikawa H, Hamamura Y, Kato M (2013) Molecular phylogeny of echiuran worms (Phylum, Annelida) reveals evolutionary pattern of feeding mode and sexual dimorphism. PLoS ONE 8(2): e56809. https://doi.org/10.1371/journal.pone.0056809

Green KD, Gambi MC (1998) Capitellidae (Annelida, Polychaeta) from the Western Mediterranean Sea: A review and new records. Italian Journal of Zoology 65(1): 1–22. https://doi.org/10.1080/11250009809386722

Grube AE (1862) Noch ein Wort über die Capitellen und ihre Stelle im Systeme der Anneliden. Archiv für Naturgeschichte, Berlin 28(1): 366–378. https://www.biodiversitylibrary.org/page/7160797

Hartman O (1947) Polychaetous annelids. Part VII. Capitellidae. Allan Hancock Pacific Expeditions 10: 391–480.

Hiley AS, Mongiardino Koch N, Rouse GW (2024) Phylogenetics of Lepidonotopodini (Macellicephalinae, Polynoidae, Annelida) and comparative mitogenomics of shallow-water vs. deep-sea scaleworms (Aphroditiformia). Biology 13(12): 979. https://doi.org/10.3390/biology13120979

Huang X, Madan A (1999) CAP3: A DNA sequence assembly program. Genome Res 9: 868–877. https://doi.org/10.1101/gr.9.9.868

Huč S, Hiley AS, McCowin MF, Rouse GW (2024) A Mitogenome-Based Phylogeny of Pilargidae (Phyllodocida, Polychaeta, Annelida) and Evaluation of the Position of *Antonbruunia*. Diversity 16(3): 134. https://doi.org/10.3390/d16030134

Hutchings P (1998) Biodiversity and functioning of polychaetes in benthic sediments. Biodiversity and Conservation 7: 1133–1145. https://doi.org/10.1023/A:1008871430178

Jennings RM, Halanych KM (2005) Mitochondrial genomes of *Clymenella torquata* (Maldanidae) and *Riftia pachyptila* (Siboglinidae): evidence for conserved gene order in Annelida. Molecular Biology and Evolution 22: 210–222. https://doi.org/10.1093/molbev/msi008

Jin JJ, Yu WB, Yang JB, Song Y, dePamphilis CW, Yi TS, Li DZ (2020) GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. Genome Biology 21: 241. https://doi.org/10.1186/s13059-020-02154-5

Katoh K, Rozewicki J, Yamada KD (2017) MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Briefings in Bioinformatics 20: 1160–1166. https://doi.org/10.1093/bib/bbx108

Kobayashi G (2023) Buried treasure in a public repository: Mining mitochondrial genes of 32 annelid species from sequence reads depos-

ited in the Sequence Read Archive (SRA). PeerJ 11: e16446. https://doi.org/10.7717/peerj.16446

Kobayashi G, Itoh H, Kojima S (2022a) Mitogenome of a stink worm (Annelida, Travisiidae) includes degenerate group II intron that is also found in five congeneric species. Scientific reports 12(1): 4449. https://doi.org/10.1038/s41598-022-08103-5

Kobayashi G, Itoh H, Nakajima N (2022b) First mitochondrial genomes of Capitellidae and Opheliidae (Annelida) and their phylogenetic placement. Mitochondrial DNA Part B 7(4): 577–579. https://doi.org/10.1080/23802359.2022.2056537

Lanfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B (2017) PartitionFinder 2: New Methods for Selecting Partitioned Models of Evolution for Molecular and Morphological Phylogenetic Analyses. Molecular biology and evolution 34(3): 772–773. https://doi.org/10.1093/molbev/msw260

Li Y, Kocot KM, Schander C, Santos SR, Thornhill DJ, Halanych KM (2015) Mitogenomics reveals phylogeny and repeated motifs in control regions of the deep-sea family Siboglinidae (Annelida). Molecular phylogenetics and evolution 85: 221–229. https://doi.org/10.1016/j.ympev.2015.02.008

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25(11): 1451–1452. https://doi.org/10.1093/bioinformatics/btp187

Magalhães WF, Blake JA (2022) Capitellidae Grube, 1862. In: Purschke G, Böggemann M, Westheide W (Eds) Handbook of zoology, Annelida, Pleistoannelida, Errantia II. Walter de Gruyter GmbH and Co KG. Berlin, Germany. https://doi.org/10.1515/9783110647167

Moore WS (1995) Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees. Evolution 49(4): 718–726. https://doi.org/10.2307/2410325

Nakaoka M, Mukai H, Chunhabundit S (2002) Impacts of dugong foraging on benthic animal communities in a Thailand seagrass bed. Ecological Research 17: 625–638. https://doi.org/10.1046/j.1440-1703.2002.00520.x

Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ (2015) IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Molecular Biology and Evolution 32(1): 268–274. https://doi.org/10.1093/molbev/msu300

Pamungkas J, Glasby CJ, Read GB, Wilson SP, Costello MJ (2019) Progress and perspectives in the discovery of polychaete worms (Annelida) of the world. Helgoland Marine Research 73(1): 1–10. https://doi.org/10.1186/s10152-019-0524-z

Perna NT, Kocher TD (1995) Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. Journal of Molecular Evolution 41: 353–358. https://doi.org/10.1007/BF01215182

Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, Finn RD (2018) HMMER web server: 2018 update. Nucleic acids research 46(W1): W200–W204. https://doi.org/10.1093/nar/gky448

Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korbeynikov A (2020) Using SPAdes de novo assembler. Current Protocols in Bioinformatics 70: e102. https://doi.org/10.1002/cpbi.102

Read G, Fauchald K (2024) World Polychaeta Database. Capitellidae Grube, 1862. [Accessed through: World Register of Marine Species] https://www.marinespecies.org/aphia.php?p=taxdetails&id=921 [on 2024-11-12]

Richter S, Schwarz F, Hering L, Böggemann M, Bleidorn C (2015) The utility of genome skimming for phylogenomic analyses as demonstrated for glycerid relationships (Annelida, Glyceridae). Genome Biology and Evolution 7(12): 3443–3462. https://doi.org/10.1093/gbe/evv224

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2012) MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. Systematic Biology 61(3): 539–542. https://doi.org/10.1093/sysbio/sys029

Rouse G, Pleijel F, Tilic E (2022) Annelida. Oxford University Press. Oxford. https://doi.org/10.1093/oso/9780199692309.001.0001

Rousset V, Pleijel F, Rouse GW, Erséus C, Siddall ME (2007) A molecular phylogeny of annelids. Cladistics 23(1): 41–63. https://doi.org/10.1111/j.1096-0031.2006.00128.x

Seixas VC, Russo CAM, Paiva PC (2017) Mitochondrial genome of the Christmas tree worm *Spirobranchus giganteus* (Annelida: Serpulidae) reveals a high substitution rate among annelids. Gene 605: 43–53. https://doi.org/10.1016/j.gene.2016.12.024

Shao R, Dowton M, Murrell A, Barker SC (2003) Rates of gene rearrangement and nucleotide substitution are correlated in the mitochondrial genomes of insects. Molecular Biology and Evolution 20(10): 1612–1619. https://doi.org/10.1093/molbev/msg176

Struck TH, Purschke G (2005) The sister group relationship of Aeolosomatidae and Potamodrilidae (Annelida: "Polychaeta")-a molecular phylogenetic approach based on 18S rDNA and cytochrome oxidase I. Zoologischer Anzeiger-A Journal of Comparative Zoology 243(4): 281–293. https://doi.org/10.1016/j.jcz.2005.01.001

Struck TH, Golombek A, Hoesel C, Dimitrov D, Elgetany AH (2023) Mitochondrial genome evolution in Annelida—a systematic study on conservative and variable gene orders and the factors influencing its evolution. Systematic Biology 72(4): 925–945. https://doi.org/10.1093/sysbio/syad023

Sun Y, Daffe G, Zhang Y, Pons J, Qiu JW, Kupriyanova EK (2021) Another blow to the conserved gene order in Annelida: Evidence from mitochondrial genomes of the calcareous tubeworm genus *Hydroides*. Molecular Phylogenetics and Evolution 160: 107124. https://doi.org/10.1016/j.ympev.2021.107124

Talavera G, Castresana J (2007) Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Syst. Biol 56: 564–577. https://doi.org/10.1080/10635150701472164

Tamura K, Stecher G, Kumar S (2021) MEGA11: Molecular evolutionary genetics analysis version 11. Molecular Biology and Evolution 38(7): 3022–3027. https://doi.org/10.1093/molbev/msab120

Tempestini A, Massamba-N'Siala G, Vermandele F, Beaudreau N, Mortz M, Dufresne F, Calosi P (2020) Extensive gene rearrangements in the mitogenomes of congeneric annelid species and insights on the evolutionary history of the genus *Ophryotrocha*. BMC genomics 21(1): 815. https://doi.org/10.1186/s12864-020-07176-8

Tomioka S, Kakui K, Kajihara H (2018) Molecular phylogeny of the family Capitellidae (Annelida). Zoological Science 35(5): 436–445. https://doi.org/10.2108/zs180009

Tsutsumi H (1987) Population dynamics of *Capitella capitata* (Polychaeta; Capitellidae) in an organically polluted cove. Marine Ecology Progress Series 36: 139–149. https://doi.org/10.3354/meps036139

Vallès Y, Halanych KM, Boore JL (2008) Group II introns break new boundaries: presence in a bilaterian's genome. PLoS ONE 3(1): e1488. https://doi.org/10.1371/journal.pone.0001488

Weigert A, Golombek A, Gerth M, Schwarz F, Struck TH, Bleidorn C (2016) Evolution of mitochondrial gene order in Annelida. Molecular Phylogenetics and Evolution 94: 196–206. https://doi.org/10.1016/j.ympev.2015.08.008

Wilkins D (2023) gggenes: Draw Gene Arrow Maps in 'ggplot2'. R package version 0.5.0. https://wilkox.org/gggenes/

Wu X, Su X, Wang Y, He C, Qiu JW, Zhang Y (2024) Unraveling the phylogeny of Chaetopteridae (Annelida) through mitochondrial genome analysis. Frontiers in Marine Science 11: 1382212. https://doi.org/10.3389/fmars.2024.1382212

Yang Z, Nielsen R (2000) Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. Molecular Biology and Evolution 17: 32–43. https://doi.org/10.1093/oxfordjournals.molbev.a026236

Zhang Z (2022) KaKs_Calculator 3.0: calculating selective pressure on coding and non-coding sequences. Genomics, Proteomics and Bioinformatics 20(3): 536–540. https://doi.org/10.1016/j.gpb.2021.12.002

Zhang Y, Sun J, Rouse GW, Wiklund H, Pleijel F, Watanabe HK, Chen C, Qian PY, Qiu JW (2018) Phylogeny, evolution and mitochondrial gene order rearrangement in scale worms (Aphroditiformia, Annelida). Molecular phylogenetics and evolution 125: 220–231. https://doi.org/10.1016/j.ympev.2018.04.002

Zhang D, Gao F, Jakovlić I, Zou H, Zhang J, Li WX, Wang GT (2020) PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. Molecular ecology resources 20(1): 348–355. https://doi.org/10.1111/1755-0998.13096

Zhong M, Struck TH, Halanych KM (2008) Phylogenetic information from three mitochondrial genomes of Terebelliformia (Annelida) worms and duplication of the methionine tRNA. Gene 416: 11–21. https://doi.org/10.1016/j.gene.2008.02.020

Supplementary material 1

Supplementary information

Authors: Xuechun Su, Deyuan Yang, Xiu Wu, Yanan Sun, Jian-Wen Qiu, Yanjie Zhang

Data type: xlsx

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/zse.101.144081.suppl1